AD_____

GRANT NUMBER: DAMD17-94-J-4371

TITLE: Computer Aided Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Carey E. Floyd, Jr., Ph.D.

CONTRACTING ORGANIZATION: Duke University Medical Center
Durham, NC 27710

REPORT DATE: February 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, MD 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 1996	Annual (23 Sep 94 - 22 Sep 95)

4. TITLE AND SUBTITLE
Computer Aided Breast Cancer Diagnosis

5. FUNDING NUMBERS
DAMD17-94-J-4371

6. AUTHOR(S)
Carey E. Floyd, Jr., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University Medical Center
Durham, NC 27710

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command,
Fort Detrick, Frederick, MD 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from the patients' mammographic findings and medical history.

In this progress period we have developed an ANN to predict biopsy outcome from mammographic and history findings. In the first year of the grant we have 1)published the preliminary data presented in the original proposal, 2)acquired 260 cases using the standardized BI-RADS reporting system and published this study, 3)conducted a small prospective study, 4)examined the inter- and intra-observer variability of the reporting lexicon, and 5)investigated reducing the number of active input features to the network and published this work. A related study was conducted to examine the sensitivity of the system to the techniques used for sampling the data. All of this work has been specifically directed toward the first specific aim of the proposal.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Computer Aided Diagnosis, Mammography, Breast Cancer	39
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to stay within the lines to meet ***optical scanning requirements.***

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denote public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

(X) Where copyrighted material is quoted, permission has been obtained to use such material.

( ) Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

( ) Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

( ) In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

( ) For the protection of human subjects, the investigator(s) have adhered to policies of applicable Federal Law 32 CFR 219 and 45 CFR 46.

( ) In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

Principal Investigator's Signature

4 Feb 96
Date

Table of Contents:

INTRODUCTION

The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from the patients' mammographic findings and medical history.

While mammography is a sensitive test for early diagnosis of breast cancer, 70% of the cases which are sent to biopsy are benign. We will develop a CAD system based on Artificial Neural Networks (ANNs) to predict the malignancy of breast findings (as verified by the outcome of biopsy) based upon radiologists' findings from mammograms. The strength of ANNs for this problem is their ability to learn complex relationships from examples of the data, then to generalize and accurately classify examples which the network has not seen before. This system will learn to predict malignancy by examining a large set of radiographic findings which are paired with biopsy results. The database for this learning will be representative of the patient population. Specifically we will, 1)Develop an ANN to predict biopsy outcome from mammographic and history findings. 2)Evaluate the improvement in radiologists' diagnostic performance when the computer diagnostic aid is provided. This implementation of an accurate CAD system will improve sensitivity, specificity, and consistency of breast cancer diagnosis and will provide a significant improvement in long term outcome for breast cancer patients.

BODY

In the first year of the grant we have published 3 peer-reviewed manuscripts[1-3] with two more accepted for publication[4][5]. Another manuscript has been submitted for peer review[6Tourassi].There have been four presentations with published proceedings at professional meetings [7][8][9][10]. In addition, we have applied for and received funding for supplemental funding to extend work on this grant[11]. Specifically, we have 1)published the preliminary data presented in the original proposal[1], 2)acquired 260 cases using the standardized BI-RADS reporting system and published this study[2], 3)conducted a small prospective study (accepted for publication[4]), 4)examined the inter- and intra-observer variability of the reporting lexicon (accepted for publication[5]), and 5)investigated reducing the number of active input features to the network and published this work[3][7][9]. A related study was conducted to examine the sensitivity of the system to the techniques used for sampling the data[6][8]. A genetic algorithm was developed as an alternative approach to the ANN [10]. All of this work has been specifically directed toward the first specific aim of the proposal.

In summary:

Year 1:

Peer-reviewed manuscripts published or in press:	5
Peer-reviewed manuscripts submitted:	1
Published Conference Proceedings:	4
International Meeting presentations:	7
Supplemental grants received:	1

Peer-reviewed manuscripts published or in press:

1. Floyd, C E, Jr, Lo, J Y, Yun, A J, Sullivan, D C, and Kornguth, P J. Prediction of breast cancer malignancy using an artificial neural network. *Cancer* **74**; 2944-2948; 1994.

2. Baker JA, Kornguth PJ, Lo JY, Williford ME, **Floyd CE Jr.**: Artificial Neural Network for the Prediction of Breast Cancer Using BIRADS Standardized Lexicon *Radiology* **196**;817-822; 1995.

3. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr. Computer-aided diagnosis of breast cancer: artificial neural network technique for optimized merging of mammographic features. *Academic Radiology*, 2;841-850;1995.

4 Baker JA, Kornguth PJ, Lo JY, Floyd CE Jr.: An Artificial Neural Network Approach to Improve the Quality of Breast Biopsy Recommendations *Radiology*Feb 1996.

5 Baker JA, Kornguth PK, Floyd CE Jr.: Bi-rads Standardized Mammography Lexicon: Observer Variability of Lesion Description (In Press: *Amer. J. Roent.* , June 1995)

Peer-reviewed manuscripts submitted:

6 Tourassi GD, Floyd CE Jr."The Effect of Data Sampling on the Performance Evaluation of Artificial Neural Networks in Medical Diagnosis". (Submitted to *Medical Decision Making* September 1995).

Published Conference Proceedings:

7 Lo JY, Grisson AT, Floyd CE Jr, Kornguth PJ. Computer-aided diagnosis of mammograms using an artificial neural network: Merging of standardized input features from the ACR lexicon. In *Proceedings of the International Society for Optical Engineering (SPIE),* **2434**;571-578; 1995.

8 Floyd CE Jr, Lo JY, Tourassi GD, Kornguth P. Computer aided diagnosis using Genetic Algorithms and Neural Networks. In *World Congress on Neural Networks,* International Neural Network Society Annual Meeting (INNS), II-863-866;1995.

9. Tourassi GD, Floyd CE Jr. Sostman HD, Coleman RE. Performance evaluation of an artificial neural network for the diagnosis of pulmonary embolism using the cross-validation, jacknife, and bootstrap methods: a comparison study . In *World Congress on Neural*

Networks, International Neural Network Society Annual Meeting (INNS),II-897-900;1995.

10. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr. Computer-aided diagnosis of mammography:Artificial neural networks for optimized merging of standardized BIRADS features. In *World Congress on Neural Networks,* International Neural Network Society Annual Meeting (INNS), II-885-888, 1995.

Meeting presentations:

11. Floyd CE Jr, Lo JY, Baker JA, Kornguth PJ: Interactive Computer aided diagnosis of Breast Cancer. *Radiology* **197P**:533, 1995.

12. Baker JA, Kornguth PJ, Floyd CE Jr: Interobserver variability in Radiologist's Use of the BI-RADS Mammography Lexicon. *Radiology* **197P**:242, 1995.

13. Baker JA, Kornguth PJ, Lo JY, Floyd CE Jr: Artificial Neural Network for the Prediction of Breast Cancer with the BI-RADS Standardized Lexicon. *Radiology* **197P**:242, 1995.

14. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr: Application of Artificial Neural Networks to the Interpretation of Mammograms on the Basis of the Radiologist's Impression and Optimized Image Features. *Radiology* **197P**:242, 1995.

15. Lo JY, Baydush AH, Baker JA, Kornguth PJ, Floyd CE Jr: Computer-Aided Diagnosis of Breast Mass Malignancy with Automated Feature Extraction and Artificial Neural Networks. *Radiology* **197P**:425, 1995.

16. Tourassi GD, Floyd CE Jr. Sostman HD, Coleman RE. Performance evaluation of an artificial neural network for the diagnosis of pulmonary embolism using the cross-validation, jacknife, and bootstrap methods: a comparison study . Presented at *World Congress on Neural Networks,* International Neural Network Society Annual Meeting (INNS),1995.

17. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr. Computer-aided diagnosis of mammography:Artificial neural networks for optimized merging of standardized BIRADS features. Presented at *World Congress on*